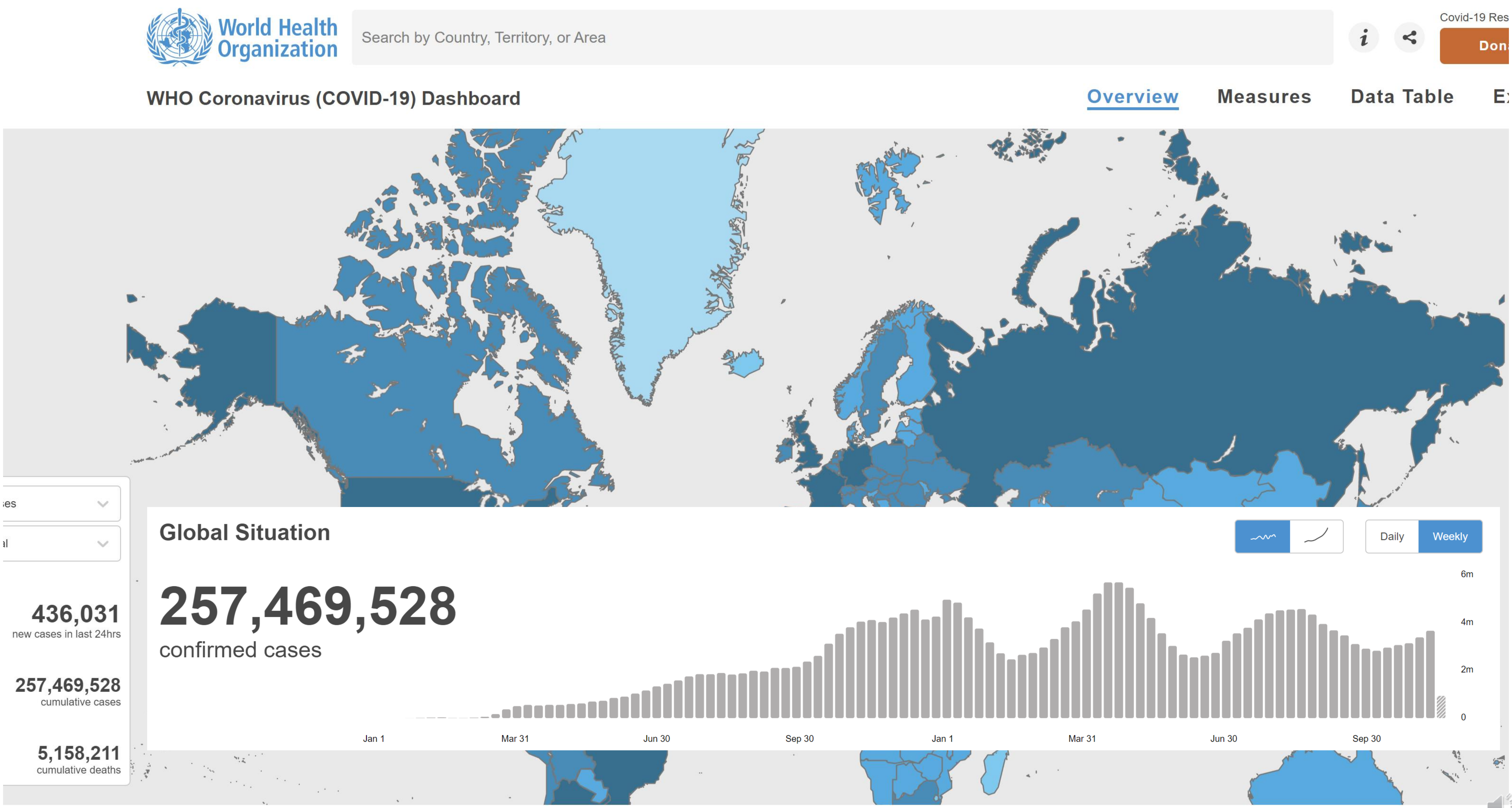
Mobility-based Individual POI Recommendation to Control the COVID-19 Spread

Tong Xia, Abhirup Ghosh

dept. Computer Science and Technology

IEEE BIG DATA 2021
Virtual Event • 15–18 December

UNIVERSITY OF CAMBRIDGE


World Health Organization
Search by Country, Territory, or Area
Covid-19 Res
Don
WHO Coronavirus (COVID-19) Dashboard
Overview
Measures
Data Table
Global Situation
Daily
Weekly
257,469,528
confirmed cases
6m
4m
2m
0
Jan 1
Mar 31
Jun 30
Sep 30
Jan 1
Mar 31
Jun 30
Sep 30
436,031
new cases in last 24hrs
257,469,528
cumulative cases
5,158,211
cumulative deaths

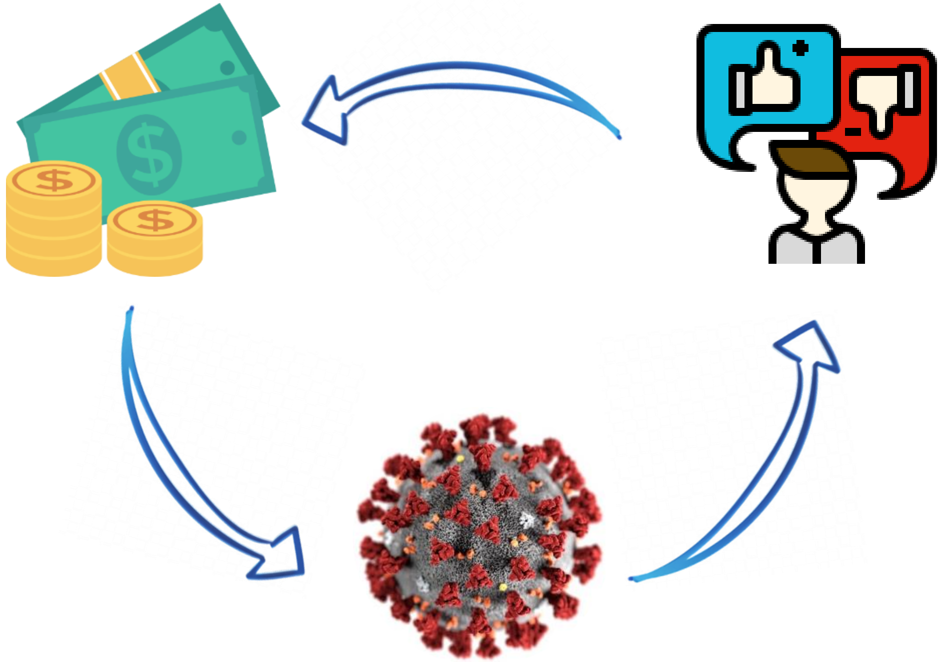
Business Survival

Personal Needs

Motivation

- **Related work for reopening:**
  - Uniform venue restrictions [2,6]
  - Super-spreader restrictions [3,4]
  - Group mobility [5,10]

Can we incorporate the insight from the structures of mobility to achieve better trade off among user preference, business development, and pandemic control?



Problem Formulation

- ***Query*** $\boldsymbol{R(v_o, t)}$: A user u will query a preferred venue v_o to be visited at time t, and this venue belongs to function category f.
- ***Recommend*** $\boldsymbol{v_r}$: A venue v_r under the same function f, within the distance threshold from v_o, and with the minimum infection risk for the whole population.

System Framework

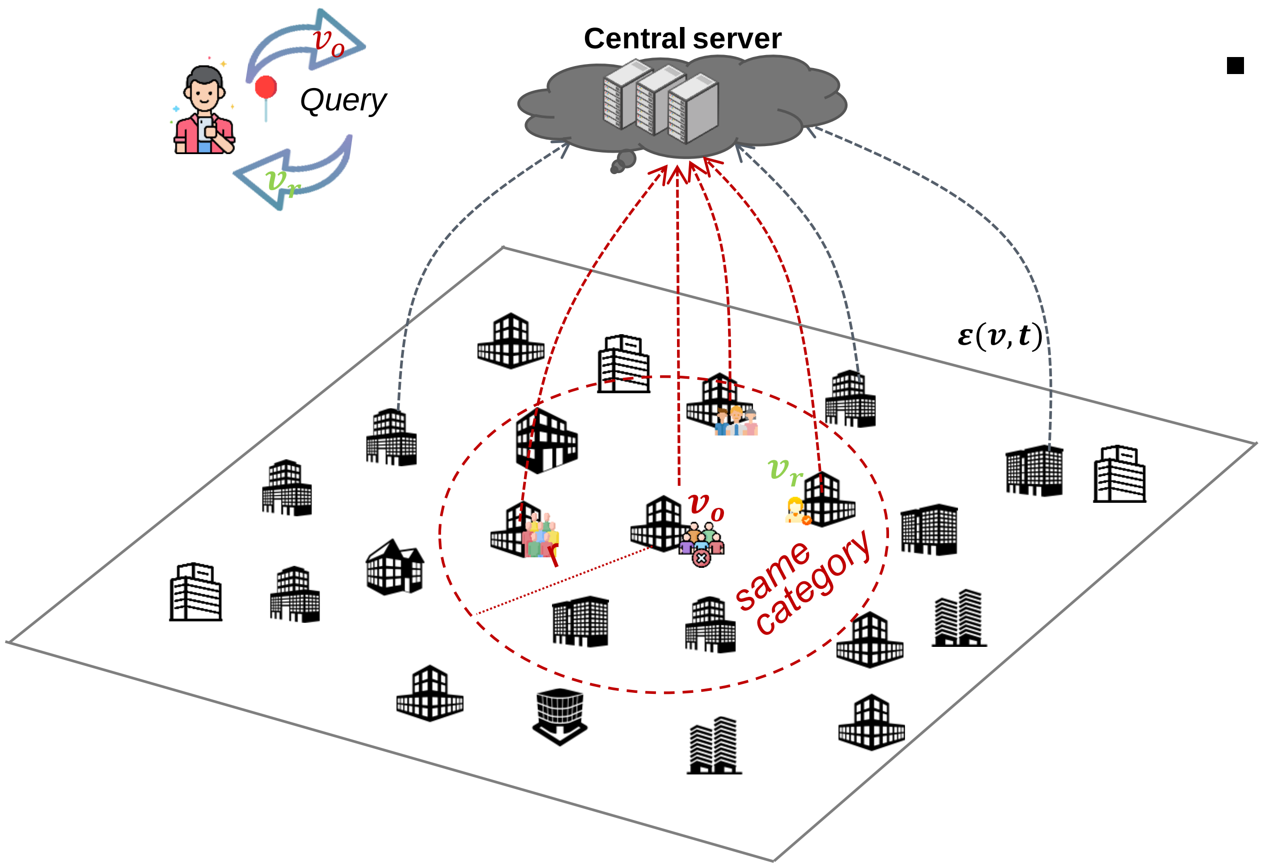


Fig. 4. An illustration of the system.

- Pipeline:
  - Maintain $\varepsilon(v,t)$ - the number of unique visitors of venue v 48 hours before t.
  - Process a coming query $R(v_o, t)$:
    - Recall POIs $v \in V_f$ and $|v - v_o| < r$.
    - Return v_r with the minimum ε.

✓ Preserving mobility demands.

✓ Do not track personally identifiable statics.

Experiments

- Real-world check-in dataset
- Agent-based COVID-19 spread model
- Extensive comparison in three cities

Real-world check-in dataset

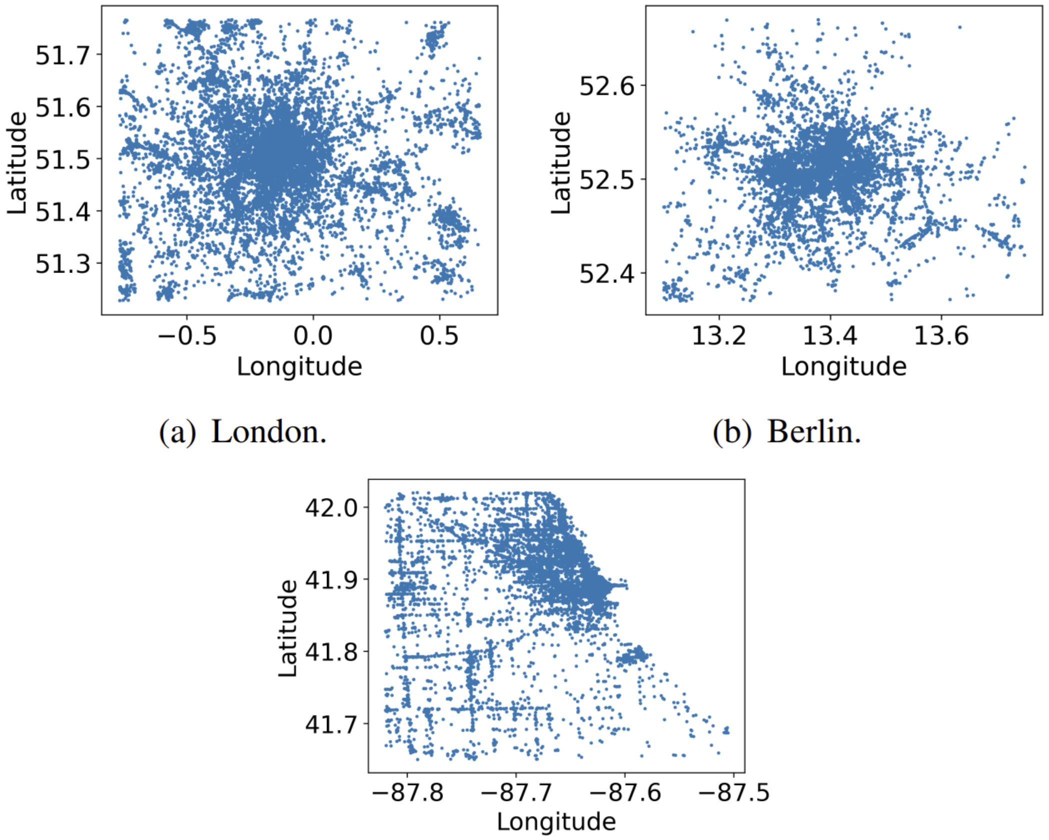
- Characteristics of the Gowalla check-in dataset

TABLE I
BASIC DATASET STATISTICS

City	#User	#Venue	#user/venue	#check-in/user
London	871	33,585	7.5	481.3
Berlin	400	12,012	7.2	274.2
Chicago	279	11,268	5.6	525.2

TABLE II
POI CATEGORIES

First Category	**Second Category**
Community	Government, Library,Worship, etc.
Entertainment	Aquarium, Stadium, Art, Theatre, Museum,etc.
Food	African, American, Asian,Bakery,BBQ,Dessert, etc.
Nightlife	Bar, Dancefloor, Microbrewery, Pub, Saloon, etc.
Outdoors	Architecture, Beach, Canal, Cemetery, etc.
Shopping	Antiques, Apparel, Bank, Bookstore, Technology, etc.
Travel	Airport, Bridge, Hotel, Subway, Train Station, etc.
Others	life severance, bicycle repair store, etc.



(a) London. (b) Berlin. (c) Chicago.

Fig. 2. Geographic distribution of venues for three cities.

Real-world check-in dataset

- Data augmentation

  1) Estimating density distributions:

     1) the *check-in temporal interval distrib*

     2) the *hourly popularity of venue* from t

     3) the *distribution of the check-in times*

  2) Sampling from density distributions:

     Through generating, we enlarge the dat

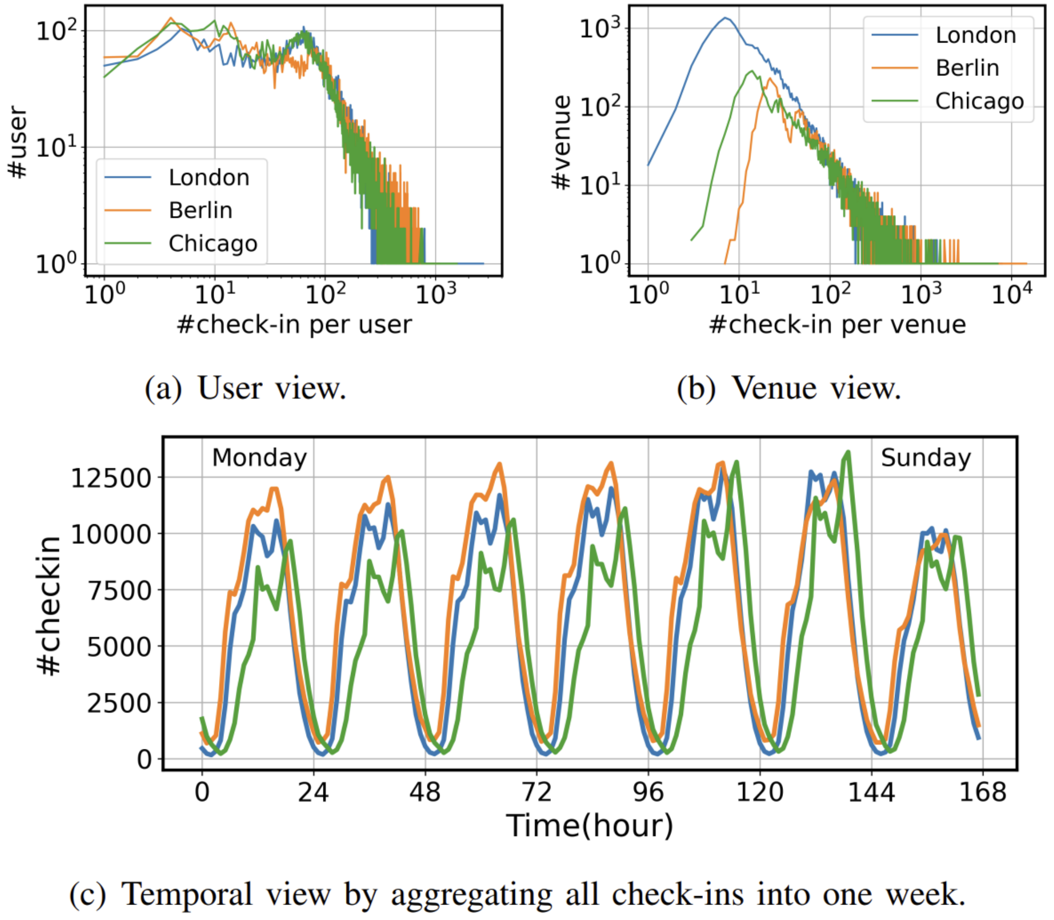


Fig. 3. Characteristics of the datasets with synthetic users for three cities.

Agent-based COVID-19 spread model

- SEIR-based model

  Susceptible-Exposed-Infectious-Removed

Infection spreads through venues

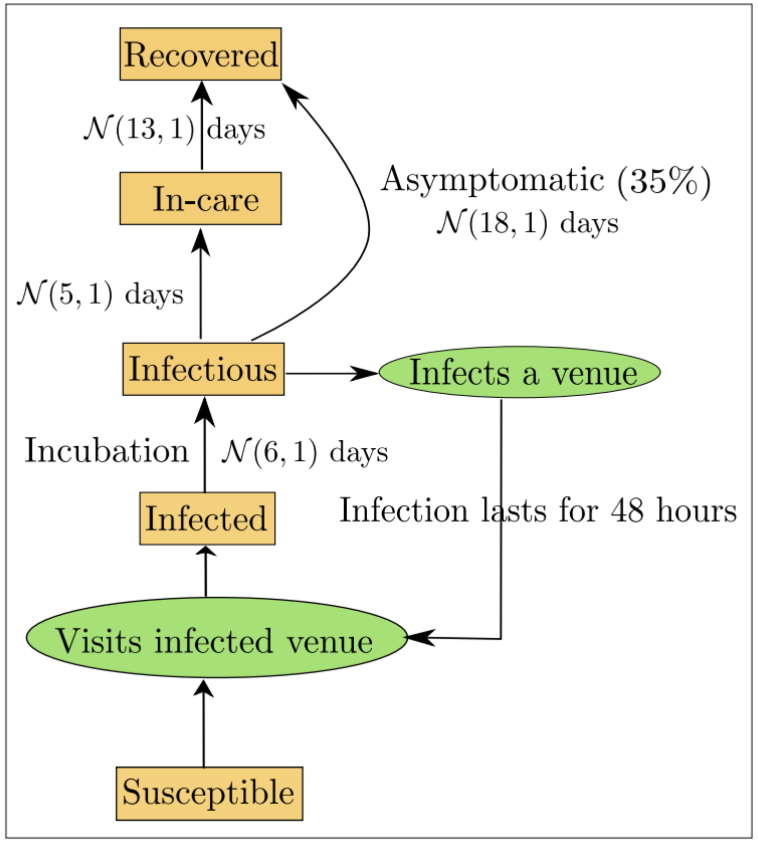
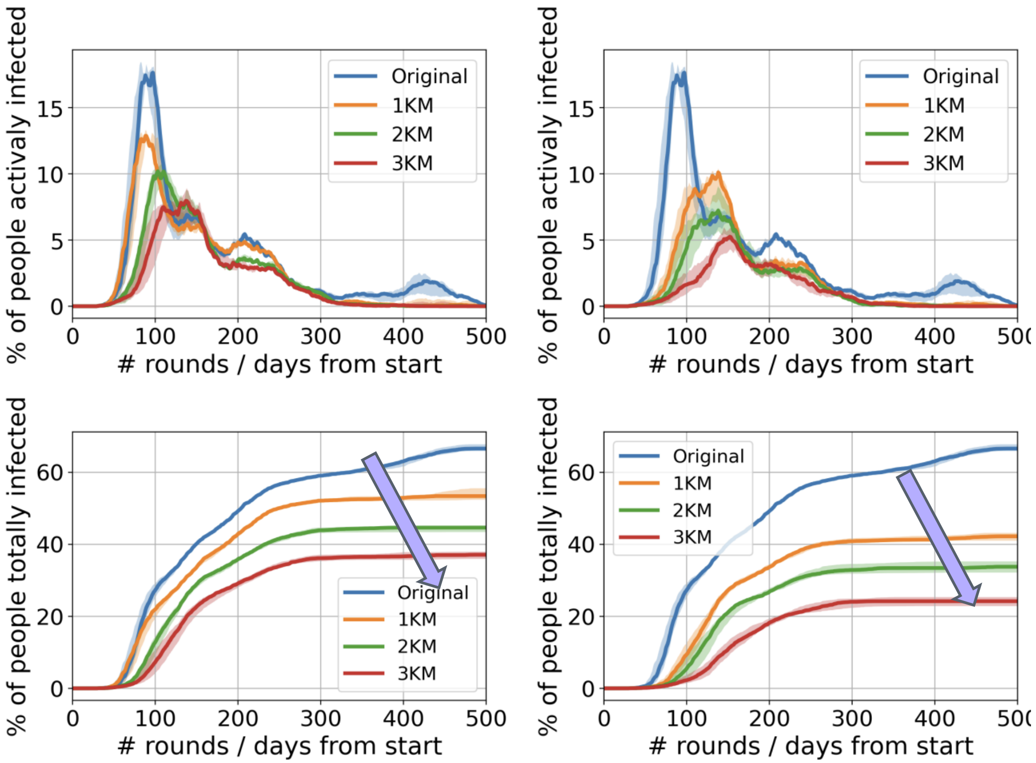


Fig. 1. The model for infection spread.

Extensive comparison in three cities

- Impact of hyper-parameters:
  - *The radius of recommendation*
  - *The granularity of POI categories*
  - *The fraction of people following the recommendation*
- Baselines:
  - *Uniformly and randomly remove x% of the check-ins.*
- Metrics:
  - *Daily new infections and the peak value.*
  - *Total infections during the simulation period.*

Results and findings -- London



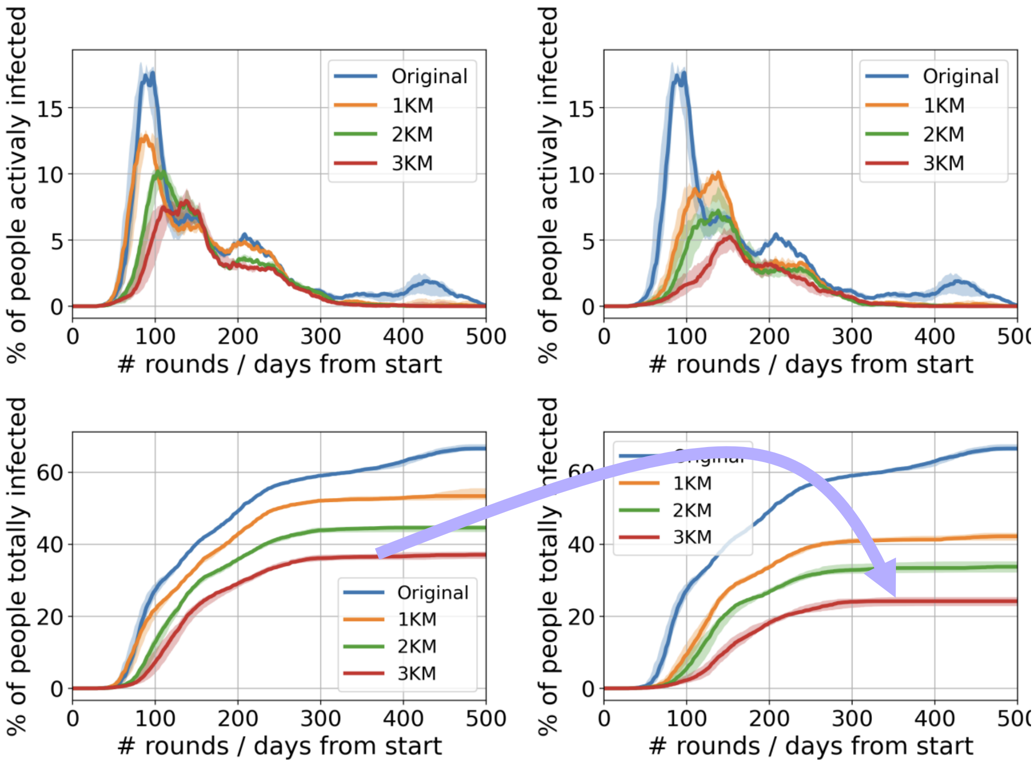
(a) Recommending venues under the same ***second category*** within various distance.

(b) Recommending venues under the same ***first category*** within various distance.

Fig. 5. Simulation results with different parameters for ***London***. The above plo of infections from the start day. The peak infection can be reduced at best to 65% following the original check-in footprints, respectively.

- *Total infection reduced by 10-20%, and peak value reduced by 5%-13% when the radium varied from 1KM to 3KM.*

Results and findings -- London



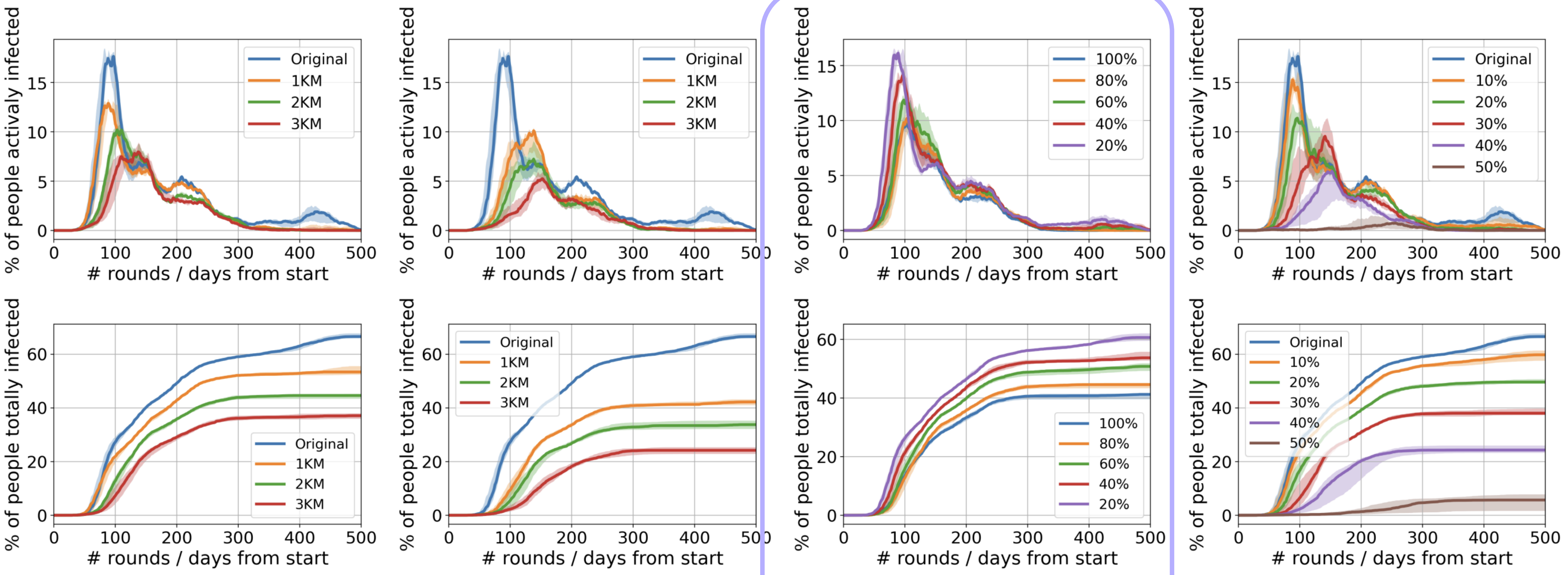
(a) Recommending venues under the same ***second category*** within various distance.

(b) Recommending venues under the same ***first category*** within various distance.

Fig. 5. Simulation results with different parameters for ***London***. The above plo of infections from the start day. The peak infection can be reduced at best to 65% following the original check-in footprints, respectively.

- *Total infection reduced by 10%-20%, and peak value reduced by 5%-13% when the radium varied from 1KM to 3KM.*
- *Total infection further reduced by 5%-20% when recommending POIs under a more general first categories*

Results and findings -- London



(a) Recommending venues under the same ***second category*** within various distance.

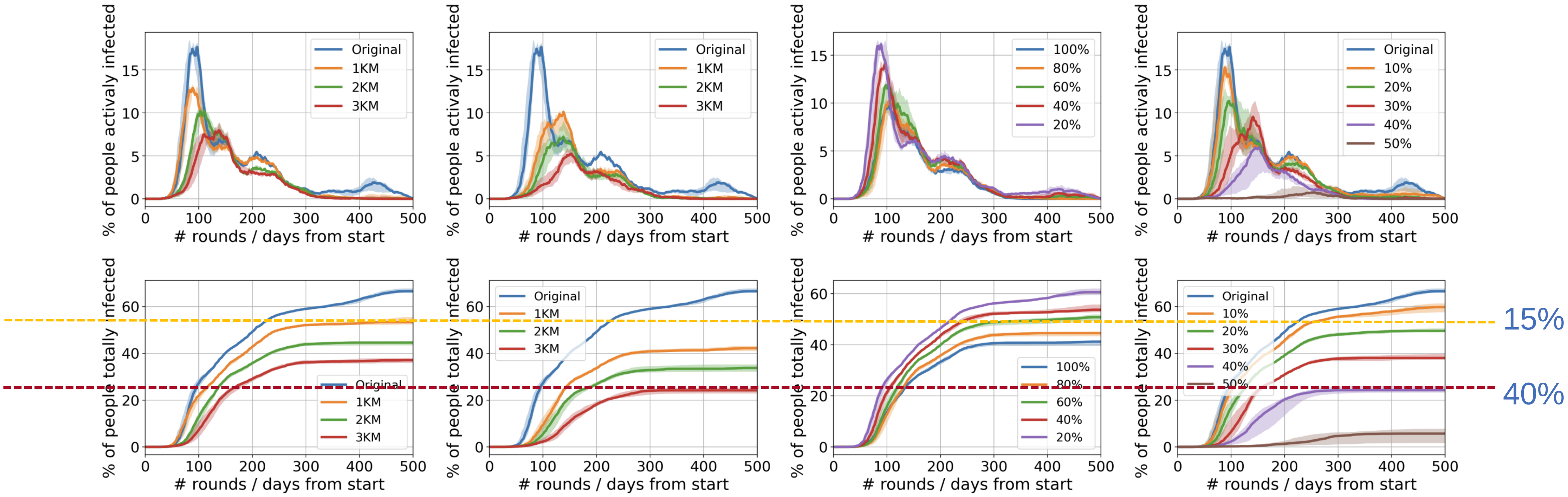
(b) Recommending venues under the same ***first category*** within various distance.

(c) A ***proportion*** of recommendations are ***followed*** (with 1KM distance threshold and under first category).

(d) ***Baseline:*** uniform mobility intervention with a certain percentage of check-ins removed.

Fig. 5. Simulation results with different parameters for ***London***. The above plots show the new infection daily, and the bottom plots present the total number of infections from the start day. The peak infection can be reduced at best to 5% and the total infections decline to as low as 22%, compared to 18% and 65% following the original check-in footprints, respectively.

Results and findings -- London



(a) Recommending venues under the same ***second category*** within various distance.

(b) Recommending venues under the same ***first category*** within various distance.

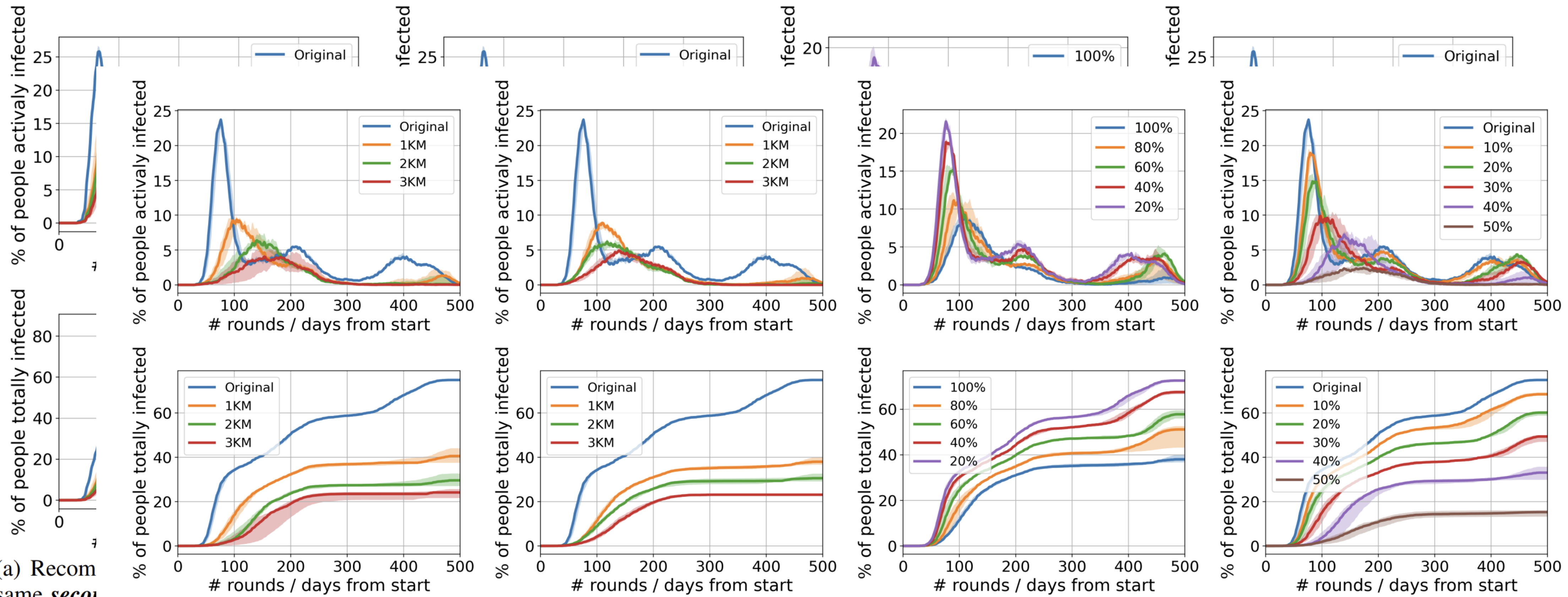
(c) A ***proportion*** of recommendations are ***followed*** (with 1KM distance threshold and under first category).

(d) ***Baseline:*** uniform mobility intervention with a certain percentage of check-ins removed.

Fig. 5. Simulation results with different parameters for ***London***. The above plots show the new infection daily, and the bottom plots present the total number of infections from the start day. The peak infection can be reduced at best to 5% and the total infections decline to as low as 22%, compared to 18% and 65% following the original check-in footprints, respectively.

This suggests that we can effectively control the disease spread while preserving all the poi visiting needs.

Results and findings -- Berlin&Chicago



(a) Recom
same *secon*
distance.

Fig. 6. Si
of infectio
84% follov

(a) Recommending venues under the same ***second category*** within various distance.

(b) Recommending venues under the same ***first category*** within various distance.

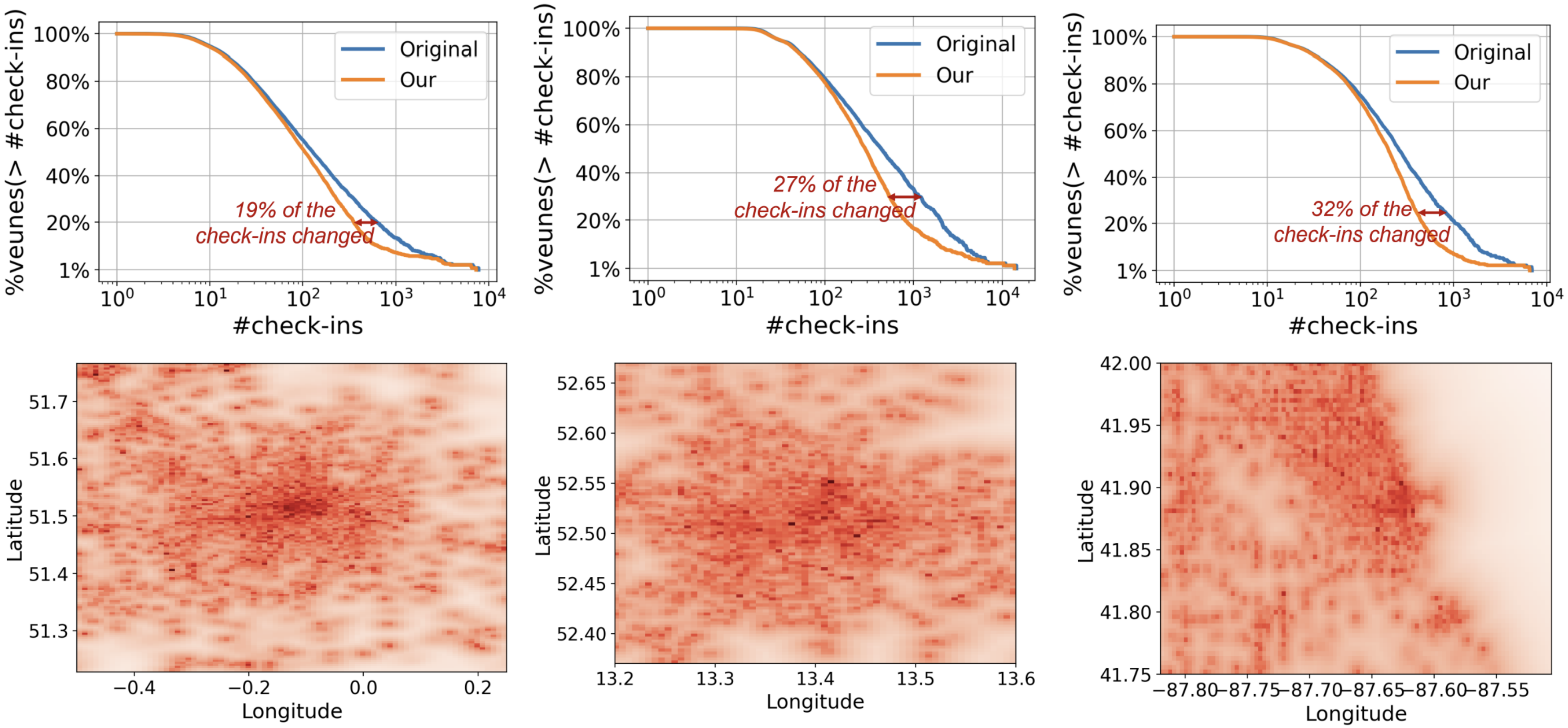
(c) A ***proportion*** of recommendations are ***followed*** (with 1KM distance threshold and under first category).

(d) ***Baseline:*** uniform mobility intervention with a certain percentage of check-ins removed.

Fig. 7. Simulation results with different parameters for ***Chicago***. The above plots show the new infection daily, and the bottom plots present the total number of infections from the start day. The peak infection can be reduced at best to 5% and the total infections decline to as low as 22%, compared to 24% and 76% following the original check-in footprints, respectively.

Results and findings

❑ Post recommendation check-in distribution



(a) ***London***. (b) ***Berlin***. (c) ***Chicago***.

Fig. 8. POI popularity distribution (the above sub-figures) and check-in density distribution (the bottom sub-figures: the darker color presents the higher density).

Conclusions

- Proposed a simple but effective method to recommend venues that can reduce the infection spread significantly while *persevering all check-in needs*.
- Our simulation results using real-world check-in datasets from three different cities verify our claims: the peak of the infection can be reduced by 5%-20%, and *the total infected population can be mitigated by at most 50%.*
- Our result is comparable to the case when 50% of check-ins are barred due to lockdown. Besides, our proposed system is also flexible and scalable to be customised based on the local COVID-19 condition.

Reference

[1] Lordan R, FitzGerald GA, Grosser T. *Reopening schools during COVID19. American Association for the Advancement of Science*; 2020.

[2] Samuel J, Rahman M, Ali G, Samuel Y, Pelaez A, et al. *Feeling like it is time to reopen now? covid-19 new normal scenarios based on reopening sentiment analytics*. arXiv preprint arXiv:200510961. 2020.

[3] Cho E, Myers SA, Leskovec J. *Friendship and mobility: user movement in location-based social networks*. In: Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining; 2011. p. 1082-90.

[4] Bonaccorsi G, Pierri F, Cinelli M, Flori A, Galeazzi A, Porcelli F, et al. *Economic and social consequences of human mobility restrictions under COVID-19. Proceedings of the National Academy of Sciences*. 2020;117(27):15530-5.

[5] Wang H, Ghosh A, Ding J, Sarkar R, Gao J. *Heterogeneous interventions reduce the spread of COVID-19 in simulations on real mobility data*. Scientific reports. 2021;11(1):1-12.

[6] Chang S, Pierson E, Koh PW, Gerardin J, Redbird B, Grusky D, et al. *Mobility network models of COVID-19 explain inequities and inform reopening*. Nature. 2021;589(7840):82-7.

Questions Are Welcomed!

Tong Xia, Abhirup Ghosh

tx229@cam.ac.uk

dept. Computer Science and Technology

IEEE BIG DATA 2021
Virtual Event • 15–18 December

UNIVERSITY OF CAMBRIDGE